Music recommendation for music learning: Hotttabs, a multimedia guitar tutor

Mathieu Barthet, Amélie Anglade, Gyorgy Fazekas, Sefki Kolozali, Robert Macrae
Centre for Digital Music
Queen Mary University of London, Mile End Road, London E1 4NS
{mathieu.barthet, amelie.anglade, gyorgy.fazekas, sefki.kolozali, robert.macrae}@eecs.qmul.ac.uk

ABSTRACT

Music recommendation systems built on top of music information retrieval (MIR) technologies are usually designed to provide new ways to discover and listen to digital music collections. However, they do not typically assist in another important aspect of musical activity, music learning. In this study we present the application Hotttabs, an online music recommendation system dedicated to guitar learning. Hotttabs makes use of The Echo Nest music platform to retrieve the latest popular or "hot" songs based on editorial, social and charts/sales criteria, and YouTube to find relevant guitar video tutorials. The audio tracks of the YouTube videos are processed with an automatic chord extraction algorithm in order to provide a visual feedback of the chord labels synchronised with the video. Guitar tablatures, a form of music notation showing instrument fingerings, are mined from the web and their chord sequences are extracted. The tablatures are then clustered based on the songs' chord sequences complexity so that guitarists can pick up those adapted to their performance skills.

Categories and Subject Descriptors

H.5.5 [**Sound and Music Computing**]: [Signal analysis, synthesis, and processing]; H.3.5 [**On-line Information Services**]: [Web-based services]; H.5.1 [**Multimedia Information Systems**]: [Video (e.g., tape, disk, DVI)]

Keywords

computer-assisted guitar tuition, automatic chord recognition, guitar tab recommendation, online music service, multimodal, hotttness measure (The Echo Nest), music video tutorial (YouTube), tag cloud, user interface

1. INTRODUCTION

The design of music recommendation systems exploiting context and/or content based information [4] has mainly been undertaken by considering music listening as the central end-user activity. Examples of such systems are the popular online music services Last.fm[1], Pandora[2], and Spotify[3], which provide new ways to experience and discover songs [1]. If, in this view, music recommendation models aim at satisfying listeners' needs and expectations, they discard other major actors of the chain of musical communication: performers. In this article, we present an online music recommendation system targeting music learning rather than music listening, therefore targeting performers rather than listeners.

Music education is one of the humanity subjects emphasised since ancient times [12]. Since the 1970s, many studies have been published in order to build computer-assisted instruction systems in various tasks of music education such as music theory, ear training, performance skills development, music composition or editing, music appreciation, musical instruments knowledge, and harmony. However, most of these systems use different approaches due to the interdisciplinary nature of the field [2, 5]. With the existing high-tech information era and the rapidly growing world wide web, it is easier to combine different musical instructions on a system to provide a good learning environment which not only integrates a variety of learning experiences for performers (e.g. textual materials, images, expert videos, and audio), but also allows individual performers to practice in less stressful conditions (e.g. the typical home setting) when compared to group-based practice [11].

Amongst musical instruments, the guitar stands out as being one of the most popular instruments in the world, with new players taking it up every day (e.g. guitar sales represent 50% of the whole musical instruments' market in France [10]). Amateur guitarists often seek new songs to play solo or with other musicians during a jam session. It is common to spend the whole time devoted to the practicing session trying to select a song adapted to one's musical skills and to find music notations in order to learn it. The proposed Hotttabs[4] application is an online guitar tuition system aimed at solving this problem by recommending popular songs to play and guiding guitarists in the learning process. Hotttabs uses a multimodal approach, relying on video tutorials, chord visualisations, and tablatures (commonly referred to

WOMRAD 2011 2nd Workshop on Music Recommendation and Discovery, colocated with ACM RecSys 2011 (Chicago, US)

[1] http://www.last.fm
[2] http://www.pandora.com
[3] https://www.spotify.com
[4] http://isophonics.net/hotttabs/

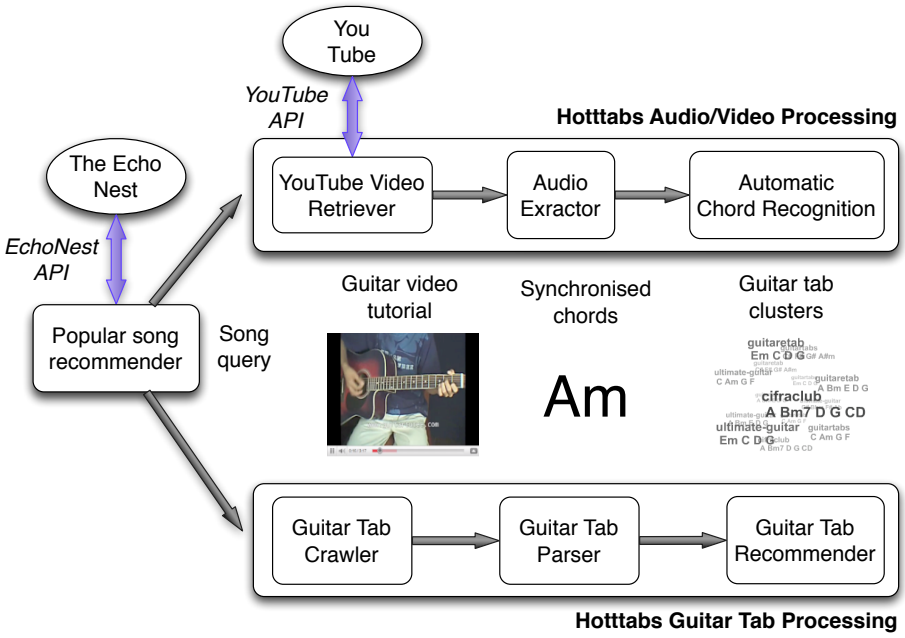

Figure 1: Hotttabs process flow chart. (API: Application Programming Interface).

as "tabs"), a form of music notation representing instrument fingering with numerical symbols rather than the musical pitches commonly used in scores.

The popularity of the guitar may be explained by several reasons: the great versatility of timbres that can be produced on acoustic or electric guitars make the instrument suitable for many different musical genres (blues, rock, reggae, jazz, classical, etc.), the simple accessibility to the instrument (guitars can be obtained at moderate costs and can be easily stored and carried away), and the possibility to learn songs regardless of prior music theory knowledge using tablatures. Since the range of pitches that can be produced on the guitar present some overlap between the various strings, notes of identical pitch can be played at several positions on the finger board (albeit with slightly different timbres due to the differences in the physical properties of the strings). One of the interests of the tablature notation is to alleviate the ambiguity on fingering by proposing an explicit solution. They can thus be considered as more effective than scores to assist beginners in guitar learning [16]. This may be one reason why guitar tabs are by far the most popular means of sharing musical instructions on the internet, largely surpassing online music sheet and MIDI score databases [13]. The Hotttabs application takes advantage of the richness of hand annotated tablatures provided on the web.

Recently, the number of guitar tuition applications for smartphones have blossomed (e.g. Killer Riffs[5], Touch Chords[6], Ultimate Guitar Tabs[7], TabToolkit[8], Rock Prodigy[9]) showing a real interest in new technologies devoted to enhancing music learning. However, most applications provide either short sections of songs, such as riffs, only the tablatures (without visual feedback showing how to play them), or predefined lists that may not contain current chart music.

In the proposed Hotttabs application, these issues are tackled by using The Echo Nest music platform[10] to retrieve the latest popular or "hot" songs based on editorial, social and charts/sales criteria, guitar video tutorials from YouTube[11], the online video sharing platform, and cutting-edge technologies in automatic chord recognition [15] and guitar tab parsing [13], to provide users with symbolic music information assisting them in the learning process. A flow chart showing the processes involved in the Hotttabs application is shown in Figure 1. The application comprises three main components: song recommendation, audio/video processing, and guitar tab processing.

The remainder of the article is organised as follows. In Section 2 we present the song recommendation method. In Section 3 we describe the audio and video processing. In section 4 the guitar tab processing is presented. Section 5 details the web application. In Section 6 we give some conclusions and perspectives on this work.

2. SONG RECOMMENDATION

The application recommends users a list of songs to practice consisting of the twenty most popular songs at the time

[5]http://itunes.apple.com/gb/app/killer-riffs/id325662214?mt=8
[6]http://itunes.apple.com/us/app/touchchords/id320070588?mt=8
[7]http://app.ultimate-guitar.com/ugt/iphone/
[8]http://agilepartners.com/apps/tabtoolkit/
[9]http://www.rockprodigy.com
[10]http://the.echonest.com/
[11]http://www.youtube.com/

of the query. These popular songs are obtained using the "hotttnesss" measure from The Echo Nest music platform. This measure which is expressed in the range [0;1] is based on editorial, social and charts/sales criteria. Editorial criteria rely on the number of reviews and blog posts that have been published about an artist in the last three weeks, providing an indicator of how much impact an artist has. Social criteria are derived from the total number of track plays the artist is receiving on a number of social media sites (for instance using statistics gathered on last.fm[12]) providing an indicator of how often people are listening to this artist. Charts/sales criteria are based on the appearance of the artist on various sales charts providing an indicator of how often people are purchasing music by this artist. A list of the twenty most popular artists is first retrieved. Then, the most popular song from each artist is selected. With such a process, the song recommender relies on a dynamic music chart directly influenced by listeners over the web, mobile or desktop applications, music consumers, and journalists.

3. AUDIO/VIDEO PROCESSING

The song learning methods underlying the application are based on a multimodal approach using audio, video, and symbolic (chord labels) feedback. We detail in this section how these modalities are exploited within the application.

3.1 YouTube guitar video tutorials retrieval

Music video tutorials offer a musical tuition alternative to music teachers since they allow one to see how a real performer plays while listening to the music. Furthermore, they often include spoken parts, giving extra information in how to perform the music or how the music is structured. YouTube provides a rich source of guitar video tutorials which are frequently updated with the latest popular songs by a large community of amateur and professional guitarists. Hotttabs filters the YouTube video database to retrieve relevant guitar video tutorials for the selected songs. To connect Hotttabs with YouTube we use Google's Data API Python client gdata[13] and request videos (YouTubeVideoQuery() function) containing the following keywords: "<song and artist> guitar chords learn".

3.2 Automatic chord recognition

Symbolic information representing music along with the video can facilitate the learning process. Furthermore in some video tutorials, the position of the player's fingers on the guitar neck cannot be seen. In order to tackle this issue, the audio tracks of the YouTube video tutorials are first extracted (using the FFmpeg converter[14]) and then processed with an automatic chord extraction algorithm. Hotttabs utilises the chord recognition algorithm described in [14] to identify the chords played by the performer and displays them on the screen synchronously with the video. This algorithm is a simplified version of the state-of-the-art chord extraction model [15] whose accuracy outperforms that obtained by typical hand annotated guitar tabs from the web [13]: the average chord accuracy (79%) obtained by the automatic method over 180 Beatles tracks is 10 percentage points higher than the chord accuracy (69%) obtained from guitar tabs.

The algorithm in [14] is implemented in the Vamp plugin[15] Chordino/NNLS Chroma[16]. A spectrally whitened log-frequency spectrum (constant-Q with three bins per semitone) is first computed. It is automatically corrected for any deviations from the standard 440 Hz tuning pitch, and an approximate semitone spaced transcription is obtained using a dictionary of notes with geometrically decaying harmonics magnitudes. The resulting semitone spectrum is multiplied with a chroma profile, and mapped to 12 bins corresponding to pitch classes. Using these features, the algorithm provides chord transcription, using a set of profiles (dictionary) to calculate frame-wise chord similarities. The resulting chord sequence is smoothed by the standard hidden Markov model (HMM)/Viterbi approach. The chord dictionary comprises the four main chord classes: major, minor, diminished, and dominant.

4. GUITAR TAB PROCESSING

One of the driving factors behind the growth in online hand annotated tabs is in the ease in which they can be produced and shared by anyone. As a consequence, these tabs do not conform to any standard format and exist in many locations on the internet. As such, we have developed methods for mining the web for these tabs and parsing them to interpret the required data.

4.1 Tab mining

The web crawler of the Hotttabs application uses 911tabs[17], a guitar tablature search engine, to access 4.7 million tabs that have already been categorised by artist, title, tablature type, and rating. Additionally, we crawled 264 common chords from Chordie[18] and Guitarsite[19] to assist in the recognition of chords when parsing tab content.

4.2 Tab parsing

To interpret the tablature text from the HTML code and the chord sequence from the tablature text, Hotttabs does the following:

- Any HTML code is stripped from the tab and "non-braking space" or "new line" tags are expanded accordingly.
- Chord definitions indicating fingerings are interpreted and added to a chord dictionary for the remainder of the tablature (e.g. "C#m: 45664-"; in this sequence, the numbers indicate the finger positions along the guitar neck for the six guitar strings ordered by decreasing pitch from left to right, and the hyphen indicates that the string must not be plucked).
- The tablature is divided up into sections based on the layout and any structural indicators (e.g. "Chorus").

[12] last.fm not only tracks listeners' musical history on their website but also when they use other services in desktop and mobile applications through what they call "scrobblers": http://www.last.fm/help/faq?category=Scrobbling

[13] http://code.google.com/p/gdata-python-client/

[14] http://ffmpeg.org/

[15] http://www.vamp-plugins.org/

[16] http://isophonics.net/nnls-chroma

[17] http://www.911tabs.com

[18] http://chordie.com

[19] http://www.guitarsite.com/chords

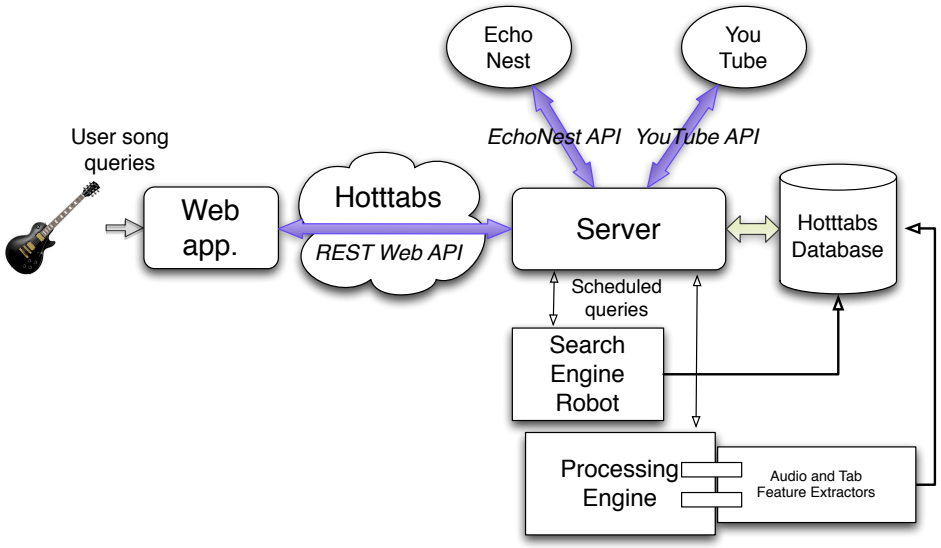

Figure 3: Hotttabs web application flow chart (app.: application; API: Application Programming Interface).

```
[Intro]
Riff1
e---------0--|--3-----3-----5-----5-----|--10-----------------------------|--------8---------
B-----3---1--|--3-----3-----7-----7-----|--12--------12--0----0---12---0---|--------10--10--7-
G------------|--4-----4-----7-----7-----|--12--------12--12---12--12---12--|--9-----9---------
D------------|--------------------------|---------------12---12--12---12--|--10--------------
A------------|--------------------------|--------------------------------|-(10)-------------
E------------|--------------------------|--------------------------------|------------------

Riff2
e--3----3----2----2----|--0----0-------(0)-|
B--3----3----3----3----|--3----3---(3)-----|
G--0----0----2----2----|--0----0-----------|
D--0----0----0----0----|--2----2----0---2--|
A--2----2----x----x----|--2----2-----------|
E--3----3----2----2----|--0----0-----------|
   G         D/F#        Em

G          D/F#        Em
Love       love        love
G          D/F#        Em
Love       love        love
D7/A       G           D7/F#    D7/E
Love       love        love
D   C   Riff3
```

Figure 2: Tab Sample. Chords Extracted: G D/F# Em G D/F# Em G D/F# Em D7/A G D7/F# D7/E D C

- Each line is scanned for any occurrence of a chord label.
- For each tab line, the tab elements are decoded accordingly.
- Any indicators of repetitions will be expanded so that "x2" will result in the current section of chords being duplicated.

An example of a tab and the chord sequence extracted can be seen in Figure 2.

4.3 Tab clustering

When learning how to play guitar, one of the difficulties lies in knowing an increasing number of chords and their relative fingerings. Thus the chord vocabulary (i.e. the set of unique chords) used in a guitar tab is of interest to the learning guitarist. Additionally, both the number of chords required to play the song and the specific chords it contains (as some chords tend to be easier to play than others) influence the guitarist when choosing a guitar tab. For any given song it is common to find several guitar tabs with chord vocabularies of varying sizes. Indeed, some simplified (e.g. to make it more accessible and easier to play) or even complexified versions (e.g. to change the style or genre of the song by varying the harmonisation) of the original song are sometimes provided on guitar tabs websites.

Thus to help the user choose between all the tabs retrieved for one seed song we further cluster the guitar tabs into three categories based on the size of their chord vocabulary: easy, medium, and difficult. To do so we rank the tabs by the number of unique chords they each contain and then divide this ranked list into three clusters. The tab clusters are then displayed as tag clouds (aggregated collections of words or phrases used to describe a document), where each tag in the cloud shows the name of the website from which the tab was extracted as well as the chord vocabulary used in the tab (see bottom of Figure 4). Therefore users can know, in one glance, which chords are required to play a given tab and how many chords it contains, without having to browse each tab individually. By clicking on an item in the tab clouds the user is redirected to the full tab in the website where it was originally published. Although the difficulty to play individual chords is not yet taken into account in the tab clustering process (which only uses the size of the chord vocabulary of the tab), displaying the chord vocabulary in the tab cloud helps users to choose the most appropriate tabs for them since they know which chords they have already learned and which chords they find difficult to play. However as most guitarists consider some specific chords to be more difficult or more tiring for the hand than others due to their fingering constraints (e.g. barre chords), we will consider including a measure of chord (fingering) difficulty into future implementations of our tab clustering algorithm.

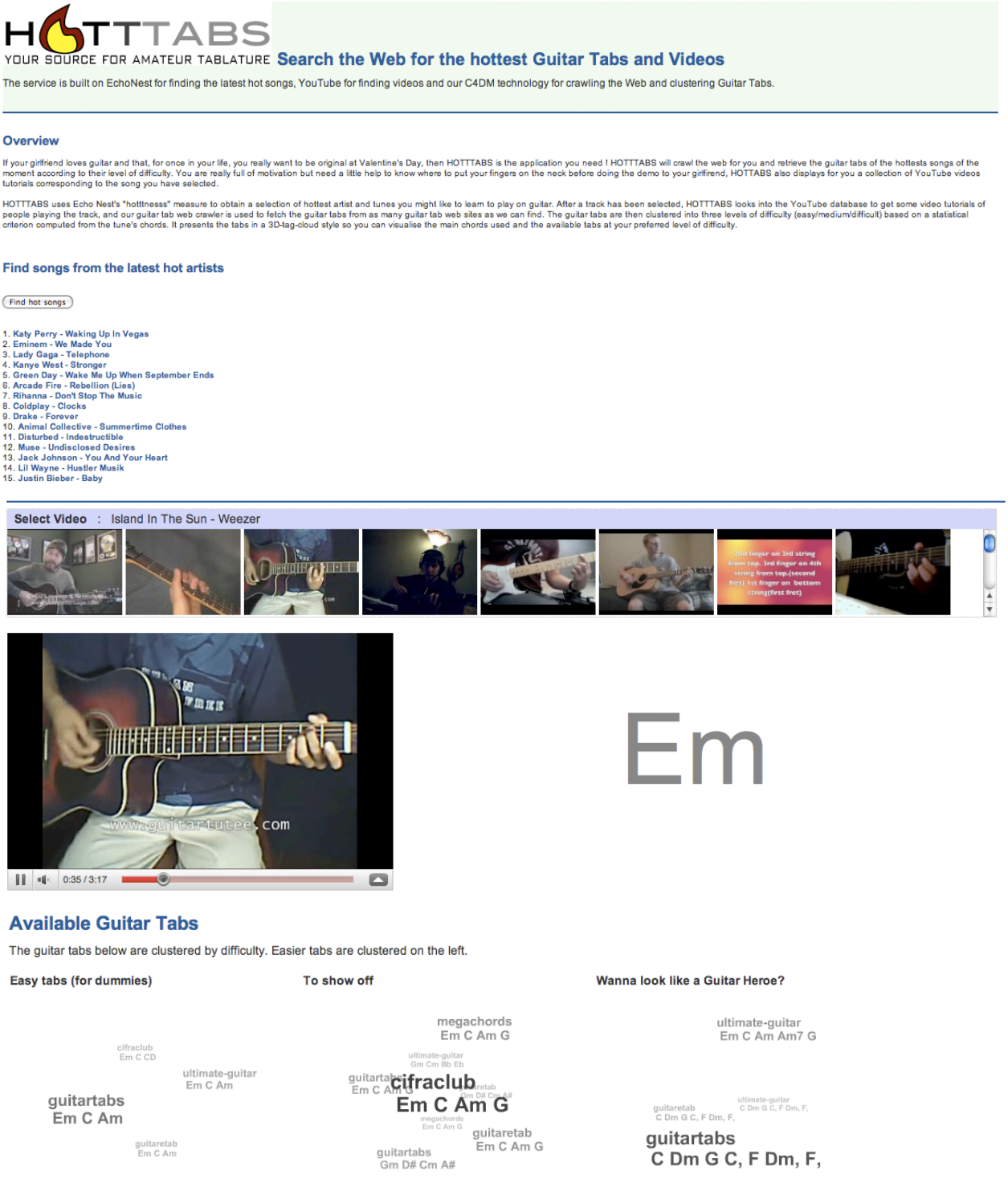

Figure 4: Screenshot of the Hotttabs web application (http://isophonics.net/hotttabs/).

5. HOTTTABS WEB APPLICATION

The Hotttabs application integrates the functionality described in the previous sections in a web-based client-server architecture.

The client runs in most popular web browsers, and provides an easy to use interface (see Figure 4). It allows the user to interact with the application and perform the following actions: *i)* query for popular songs, *ii)* retrieve a list of video tutorials and three sets of tab clusters (easy, medium, and difficult) for the selected popular song, *iii)* play a video, from a list of thumbnails, in an embedded video player, synchronised with automatically extracted chords, *iv)* select and link to a tab from the tab clusters as you would from a search engine.

In response to user interaction, the server performs the core functionality as described in section 5.2. Concerning client-server communication, Hotttabs follows the Representational State Transfer (REST) style web application design (see Figure 3). In this architecture web pages form a virtual state machine, allowing a user to progress through the application by selecting links, with each action resulting in a transition to the next state of the application by transferring a representation of that state to the user [9].

5.1 Front end

The light weight client uses a combination of standard web technologies (HTML, CSS, JavaScript) and makes use of the JQuery[20] library to dynamically load content from the server via XMLHttp requests. This content includes the list of popular songs, and the a list of video thumbnails for a selected song. We developed client-side JavaScript code which interacts with the embedded YouTube player, to display chord names next to the video. The chord sequence is requested when the user starts the video, and returned using JavaScript Object Notation with timing information, which is used to synchronise the chords with the video. The tab clusters are displayed using an adapted version of the WP-Cumulus Flash-based tag cloud plugin[21]. This plugin utilises XML data generated on the server side from the results of the tab search and tab clustering algorithm.

5.2 Back end

The server side of the Hotttabs application builds on semantic audio and web technologies outlined in [8]. The Sonic Annotator Web Application (SAWA) [7], a Python[22] framework for writing web applications involving audio analysis, is used as a basis for Hotttabs. This is extended with modules to access The Echo Nest, YouTube, and perform additional application specific functionality as shown in Figure 3.

The communication between the client and server is coordinated using the Model View Controller (MVC) architectural pattern [6]. Some important domain objects in the MVC model, as well as the Hotttabs database, are provided by the Music Ontology framework [17], such that corresponding data structures are generated from the ontology specification using the Music Ontology Python library [7]. For instance, information about popular artists and their songs (retrieved from The Echo Nest) are stored in objects and database entries corresponding to the `mo:Track`[23] and `mo:MusicArtist` concepts.

Besides user interaction, the server also performs scheduled queries for popular songs to bootstrap the database. This is necessary, since crawling for guitar tabs and the feature extraction process for chord analysis are too computationally expensive to be performed in real-time. This process uses the crawler described in section 4.1, as well as the chord extraction algorithm of [14] implemented as a Vamp audio analysis plugin [3] which can be loaded by the processing engine of SAWA.

6. CONCLUSIONS AND PERSPECTIVES

We presented Hotttabs, an online multimedia guitar tuition service comprised of the following features: (i) the recommendation of popular songs based on The Echo Nest "hotttness" measure, taking into account the artists' popularity dynamically through web data mining, (ii) the retrieval of guitar video tutorials from the YouTube database, (iii) the visual feedback of the chord labels using a content-based music information retrieval technique, and (iv) the recommendation of guitar tablatures targeting users of different levels depending on the vocabulary of chords in the selected song.

We plan to conduct a user survey in order to obtain some feedback to feed into future technical developments of the application. We also intend to model user skills and assess performances in order to adapt which music and guitar tabs are recommended, based on the users learning process. Interesting follow-ups to this work also include the development of a guitar chord fingering dictionary to display various possible chord fingerings along with the chord labels. The chord concurrence measure introduced in [13] could be used to select the most accurate guitar tabs and discard erroneous ones. Future work will also address the development of new tab clustering methods based on the chord sequence parsing, the integration of an audio/video time-stretching technique to allow for the slowing down of the video tutorials, and the synchronisation of guitar tabs and lyrics with the videos using sequence alignment.

7. ACKNOWLEDGMENTS

The authors would like to thank Matthias Mauch for his automatic chord recognition system. This work was partly funded by the musicology for the masses EPSRC project EP/I001832/1, and the Platform Grant from the Centre for Digital Music funded by the EPSRC project EP/E045235/1.

8. REFERENCES

[1] P. Aman and L. A. Liikkanen. A survey of music recommendation aids. In *Proceedings of the 1st Workshop on Music Recommendation and Discovery (WOMRAD) colocated with ACM RecSys*, pages 25–28, Barcelona, Spain, 2010.

[20] http://jquery.com/

[21] http://wordpress.org/extend/plugins/wp-cumulus/

[22] http://www.python.org

[23] In this notation, the namespace prefix **mo:** represents the URL of the Music Ontology.

[2] M. Brandao, G. Wiggins, and H. Pain. Computers in music education. In *Proceedings of the AISB'99 Symposium on Musical Creativity*, Edinburgh, UK, 1999.

[3] C. Cannam. The VAMP audio analysis plugin API: A programmer's guide. Availble online: `http://vamp-plugins.org/guide.pdf`, 2009.

[4] O. Celma. *Music Recommendation and Discovery - The Long Tail, Long Fail, and Long Play in the Digital Music Space*. Springer, 2010.

[5] Y.-K. Cliff Liao. Effects of computer-assisted instruction on students' achievement in Taiwan: A meta-analysis. *Computers & Education*, 48:216–233, 2007.

[6] E. Evans. *Domain-Driven Design: Tackling Complexity in the Heart of Software*. Addison-Wesley Professional, 2003.

[7] G. Fazekas, C. Cannam, and M. Sandler. Reusable metadata and software components for automatic audio analysis. *IEEE/ACM Joint Conference on Digital Libraries (JCDL'09) Workshop on Integrating Digital Library Content with Computational Tools and Services, Austin, Texas, USA*, 2009.

[8] G. Fazekas, Y. Raimond, K. Jakobson, and M. Sandler. An overview of Semantic Web activities in the OMRAS2 project. *Journal of New Music Research (special issue on Music Informatics and the OMRAS2 project)*, 39(4):295–311, 2010.

[9] R. T. Fielding and R. N. Taylor. Principled design of the modern Web architecture. *ACM Transactions on Internet Technology*, 2(2):115–150, 2002.

[10] IRMA. Le marché des instruments de musique: une facture bien réglée [the market of musical instruments: a well tuned invoice]. `http://www.irma.asso.fr/LE-MARCHE-DES-INSTRUMENTS-DE?xtor=EPR-75`, Last viewed in September 2011.

[11] A. LeBlanc, Y. C. Jin, M. Obert, and C. Siivola. Effect of audience on music performance anxiety. *Journal of Research in Music Education*, 45(3):480–496, 1997.

[12] S.-J. Lou, Y.-C. Guo, Y.-Z. Zhu, R.-C. Shih, and W.-Y. Dzan. Applying computer-assisted musical instruction to music appreciation course: An example with Chinese musical instruments. *The Turkish Online Journal of Educational Technology (TOJET)*, 10(1), 2011.

[13] R. Macrae and S. Dixon. Guitar tab mining, analysis and ranking. In *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR 2011)*, Miami, Florida, USA, 2011.

[14] M. Mauch and S. Dixon. Approximate note transcription for the improved identification of difficult chords. In *Proceedings of the 11th International Society for Music Information Retrieval Conference (ISMIR 2010)*, Utrecht, Netherlands, 2010.

[15] M. Mauch and S. Dixon. Simultaneous estimation of chords and musical context from audio. *IEEE Transactions on Audio, Speech, and Language Processing*, 18(6):1280–1289, 2010.

[16] A. Naofumi, A. Yoshiano, and Y. Tsuyoshi. A study of automated tablature generation for assisting learning playing the guitar. Technical Report 189, IEIC Technical Report (Institute of Electronics, Information and Communication Engineers), Japan, 2000.

[17] Y. Raimond, S. Abdallah, M. Sandler, and G. Frederick. The music ontology. In *Proceedings of the 7th International Society for Music Information Retrieval Conference (ISMIR 2007)*, Vienna, Austria, 2007.